CN 53-1189/P ISSN 1672-7673

Hα 全日面云污染实时识别和修复系统*

张能维，杨云飞，李冉阳，季凯帆

（昆明理工大学云南省计算机技术应用重点实验室，云南 昆明 650500）

摘要：高空云层导致所观测的Hα全日面像上覆盖有一层云污染，使得图像上的太阳活动细节变得模糊不清。为了能够实时探测云污染，并及时显示修复后的图像，采用图形处理器技术实现了一个Hα全日面云污染实时识别和修复系统。该系统主要在统一计算设备架构(Compute Unified Device Architecture, CUDA)环境下利用图形处理器并行实现：(1)二值化图像椭圆长短轴比值法识别重度云污染图像；(2)临边昏暗曲线中心对称法识别可修复云污染图像；(3)频域巴特沃斯低通滤波法去除云污染。通过对系统中各运算在图形处理器中花费的时间进行详细测量，发现傅里叶正反变换和频域滤波占用了图形处理器总处理时间的52.9%，是系统中最耗时的。然而，相对于1 min的观测时间间隔，约0.7 s的总处理时间可以满足实时显示的需要。另外，通过对修复后的图像做质量评价，验证了所采用的修复算法可以有效地去除云污染，并对太阳活动细节影响较小。最后讨论了系统存在的不足和需要进行的改进。

关键词：云污染；并行计算；识别和修复；巴特沃斯低通滤波器

中图分类号：P111；TP274 **文献标识码**：A **文章编号**：1672-7673(2016)02-0242-08

Hα全日面观测为监测太阳色球层活动现象提供大量的数据。全球拥有众多的Hα全日面观测站，如全球日震网(Global Oscillation Network Group, GONG)拥有全球6个观测点[1]，国内的北京怀柔观测基地等等。虽然各站点在选址时进行了严格的筛选，但是实际观测中会出现大量的多云天气[2]。在这样的天气观测，经常出现太阳局部被云覆盖的情况，导致图像细节模糊。图1显示了Cerro Tololo美洲际天文台在2015年4月20日观测的3张图像，可以看出后两张图像存在不同程度的云污染现象。

云覆盖是导致全日面像畸变的一种情况。为了检测和校正各种畸变，前人提出了一些关于Hα全日面像的质量评价和修复算法。如文[3]采用双尺度滤波归一化图像强度；文[4]根据图像临边昏暗曲线的峰度判断图像质量；文[5]根据最小相关系数法判断图像是否有云污染；文[6-7]采用大尺度中值滤波修复畸变图像和去除云污染，文[5]使用多尺度形态学滤波法修复云污染图像等。这些方法主要为了在后续的太阳物理研究中，能在观测图像上更好地切割、识别和测量各种太阳活动特征结构，如耀斑、暗条等等。因此，并不涉及执行效率和运行速度的问题。

如果能将这些方法用于实时探测是否有云污染，并实时显示修复后的图像，将非常有利于观测的进行和观测质量的判断。但这必然要求相应的优化算法和计算机的快速处理能力。为此基于文[5]的处理方法，采用图形处理器技术实现了一个实时全日面像云污染的识别和修复系统。在硬件环境下，处理一张云污染图像消耗的总时间约为0.7 s。根据图形处理器技术的特点，采用频域巴特沃斯滤波器，并使用图像相似度算法对无云图像和云污染修复后的图像进行局部相似度检验。结果显示，在满足处理速度的同时，这一滤波算法可以在很大程度上保留太阳的活动细节，满足观测实时显示的需要。

* 基金项目：国家自然科学基金（11303011，11263004，11463003，11163004，U1231205）资助.
收稿日期：2015-06-18；修订日期：2015-07-13
作者简介：张能维，男，硕士. 研究方向：天文图像处理. Email：281689743@qq.com
通讯作者：杨云飞，女，副教授. 研究方向：计算机应用与天文图像处理. Email：flyingwithcloud@sina.com

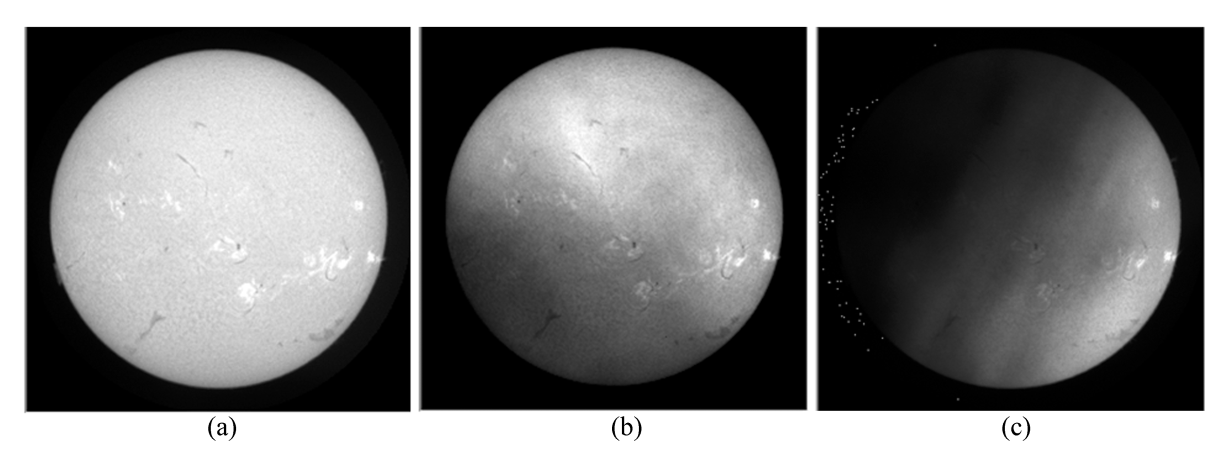


图 1 （a）、（b）、（c）3 幅图像为 Cerro Tololo 美洲际天文台分别于当地时间 2015 年 4 月 20 日（16∶07、16∶08、20∶17）拍摄

Fig. 1 （a），（b）and（c）are three Hα full-disk solar images，which were observed by the GONG Hα instrument of the CTIO on April 20，2015

1　算法原理分析

对云污染图像的处理包括两步：云污染的检测和云污染的修复。其中检测包括：(1)判断全日面像是否为一个圆，否则认为是重度云污染的图像而放弃修复；(2)判断全日面像的临边昏暗曲线是否圆对称，如果不对称，说明有云污染。云污染的修复采用文[5]提出的模板法，即通过和一个标准全日面像模板进行对比，并通过滤波得到云层透过率，从而修复被污染的图像。

1.1　重度云污染图像的判断

正常情况下的全日面像其太阳圆面与天空背景的强度值相差很大，可以用一个高于天空背景的阈值对其进行二值化。当云污染过重时，日面像就会残缺。图 2 为图 1(c)二值化后的图像。

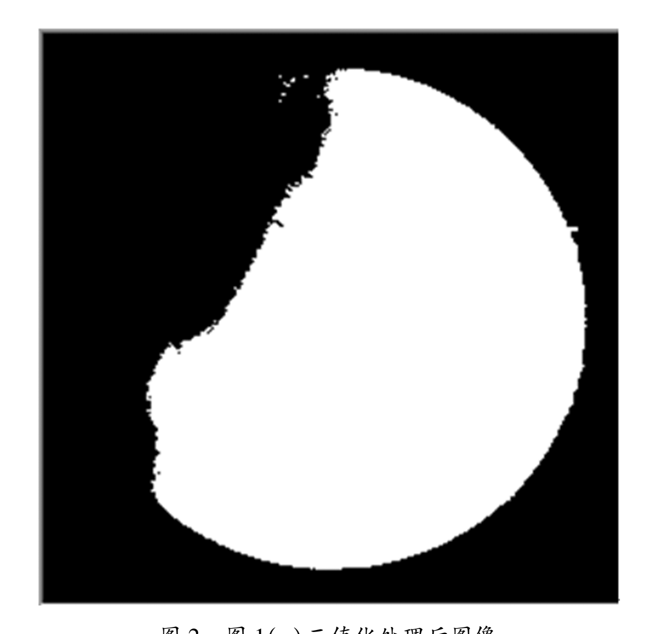


图 2　图 1(c)二值化处理后图像

Fig. 2　Binary image processing result of Fig. 1(c)

因此，可先将二值化图像假定为一个椭圆，通过计算其长轴与短轴的比值 E 判断是否残缺。对良好观测条件下的 GONG 图像统计表明，其 E 值为 1.0010 ± 0.0015。而图 2 的 E 值为 1.938，可判定为重度污染的图像。在实际系统中，凡 E 值高于 1.1 的都归为重度污染的图像。

1.2　无云污染图像的判断

对于二值化后为一个圆的图像，多数没有云污染。根据文[5]的方法，可通过判断临边昏暗曲线是否中心对称判断是否有云污染存在。主要步骤包括：(1)将太阳图像中心化并将其从直角坐标系转到极坐标系；(2)获取极坐标图像等分 4 部分中值临边昏暗曲线；(3)4 条中值临边昏暗曲线两两求相关系数，比较得到最小相关系数。

如果无云污染，这个系数接近 1，反映了临边昏暗曲线的圆对称性。当有云污染时，这个系数会变小。在系统中采用的阈值为 0.95。图 3 显示了图 1(b)的 4 条临边昏暗曲线，其最小相关系数为 0.32，而相应的图 1(a)为 0.99。据此判定图 1 中(a)不存在云污染，(b)存在可修复的云污染。

1.3　云污染的修复

云污染全日面像的成像模型可简化为以下公式：

$$T(x, y) = c(x, y)t(x, y),$$

其中，$c(x, y)$为云层透射率；$t(x, y)$为没有云污染的全日面像；$T(x, y)$为实际观测的云污染全日面像。修复原理可简单表示成$t(x, y) = T(x, y)/c(x, y)$。修复的步骤为：(1)选择一张无云图像(如图1(a))处理得到其标准临边昏暗轮廓(图4(a))作为模板；(2)用有云污染的图像(图1(b))除标准临边昏暗轮廓模板，得到含有太阳活动细节的云层透射率(图4(b))；(3)将图4(b)转换到频域并利用巴特沃斯滤波器对其进行低通滤波，扣除太阳活动细节的影响，得到云层透射率(图4(c))；(4)将原始的云污染图像除以云层透射率得到最终去除云污染的图像(图4(d))。

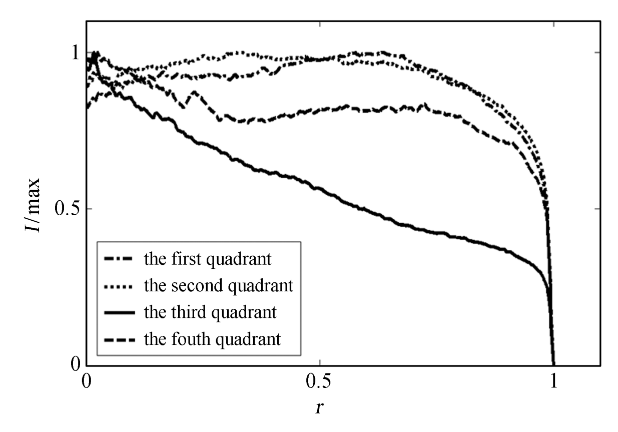


图3 图1(b)对应的四个象限的中值临边昏暗曲线

Fig. 3 The median limb-darkening profiles of four quadrants in azimuth of Fig. 1(b)

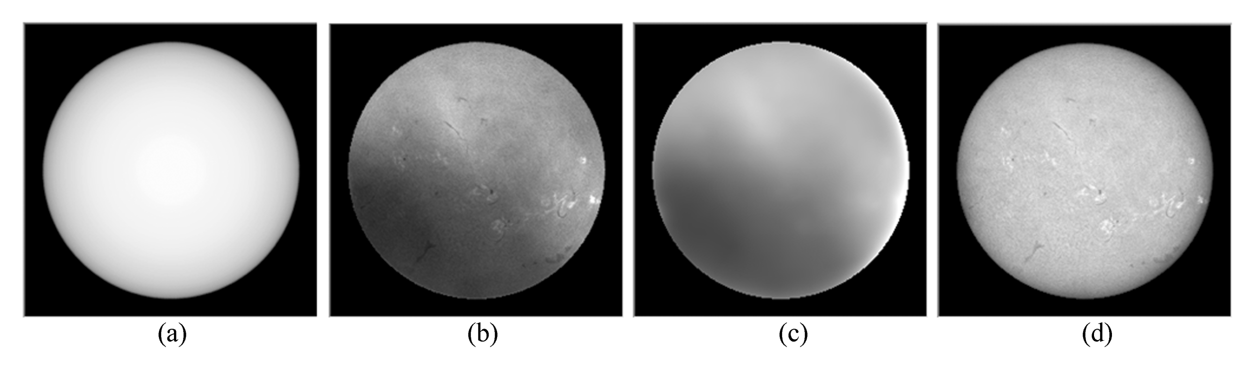


图4 (a)由图1(a)所得标准临边昏暗轮廓；(b)含太阳活动细节的云层透射率；(c)去除太阳活动细节的云层透射率；(d)去除云污染的图像

Fig. 4 (a) The standard profile of Fig. 1(a); (b) The clouds transmittance with solar features; (c) The clouds transmittance without solar features; (d) The solar image without cloud pollution

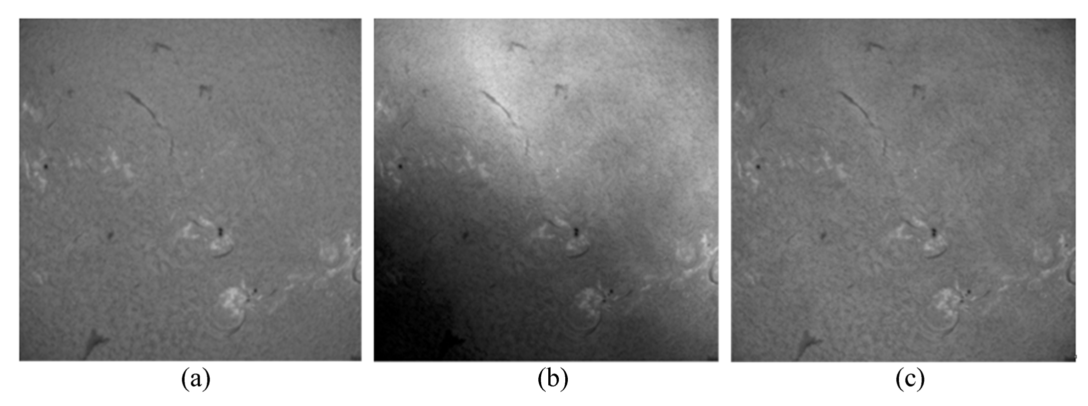
在第(3)步发现，图4(b)中的太阳活动细节在频域图中主要分布于以中心为原点，半径为D_0的圆外区域。因此在巴特沃斯滤波器中，只需选择合适的半径D_0(截止频率)将其圆外的高频部分滤除，即可在滤波后得到扣除太阳活动细节的云层透射率。经过统计不同时段的云污染图像，将系统中截止频率取为图像尺寸的0.003倍。

为充分发挥图形处理器的并行处理优势，采用频域滤波法。为此，采用图像结构相似度(Structural Similarity Index Measurement, SSIM)[8]算法证实修复的可靠性。选两幅拍摄时间仅隔一分钟的图像，第1幅是干净的图像，第2幅为云污染的图像，对第2幅图像进行修复得到第3幅图像。图5显示了这3幅图像的局部区域。计算得到无云图像(图5(a))与有云图像(图5(b))的图像结构相似度为0.2997，无云图像(图5(a))与去除云图像(图5(c))的图像结构相似度为0.7558。可以看出去除云污染之后图像的结构相似性明显高于去除云污染之前的图像，说明频域滤波也可以相当程度地去除云污染，并保留了太阳的活动细节。

2 云污染图像识别和恢复系统的图形处理器实现

图形处理器的设计初衷是为了将计算机中关于图形图像的处理和显示从中央处理器中独立出来。由于图形处理器采用了单指令多数据流[9]的数据操作方式，因而并行计算能力很强。为了将图形处理器的多线程并行能力运用到通用计算上来，NVIDIA公司推出了统一计算设备架构这一编程模型[10]。全日面像云检测和修复算法具备数据计算密度高、可并行性强的条件，因而十分适合用图形

处理器处理。并且 CUDA 中有很多现成的函数库，如 CUBLAS 库中的大量针对矩阵的基本运算函数以及 CUFFT 库中的傅里叶正反变换函数，可以极大提高程序的运行效率。例如，文[11]利用图形处理器技术在国家天文台怀柔观测站实现一个实时深积分磁场观测系统，文[12]利用图形处理器技术实现了抚仙湖 1 m 真空太阳望远镜 level 1 级图像选帧。



(a)　　(b)　　(c)

图 5　(a)、(b)、(c)分别为图 1(a)、图 1(b)、图 3(d)自中心截取的 1 200 × 1 200 像素局部图

Fig. 5　(a), (b) and (c) are three 1200 × 1200 partial image, which are intercepted from the center of Fig. 1(a), Fig. 1(b) and Fig. 3(d)

图 6 展示了系统流程图。该系统首先要显示实时处理界面，然后循环检测是否有更新的图像。如果发现更新图像先显示原图，再做云图像检测和修复处理，最后把处理后的图像显示在界面上与原图像对比。如果没有检测到更新图像，则界面一直显示之前图像的处理结果。

其中中央处理器所要完成的主要任务包括：(1)将原始 FITS 格式图像数据从硬盘读入中央处理器内存并显示在界面上；(2)将图像数据从中央处理器内存复制到图形处理器显存；(3)将处理后得到的数据从图形处理器显存复制回中央处理器内存并显示。

图形处理器执行系统的核心任务——云检测和云去除。将图形处理器中计算任务大致分为两类：一类是通过直接调用 CUDA 库函数实现。另一类是自行编写的利用线程并行实现的函数，其执行的主要思想为：(1)根据数矩阵点的总数分配总线程数，通过线程号索引让矩阵中每个点与线程一一对应；(2)根据判断条件将每个线程要完成的计算任务分类，同一判断条件下所有线程的计算任务相同；(3)所有线程并行执行各自分配的计算任务。

在检测重度云污染图像环节，首先并行实现图像二值化处理。椭圆长轴与短轴并行求取方法为：(1)并行求取椭圆重心位置；(2)并行得到椭圆上各点与其重心的协方差。最后将所得协方差元素代入公式即可得到椭圆长短轴之比。据此如果判断图像云污染过重，则丢弃，否则继续下面步骤。

接下来要检测图像是否存在可修复的云污染。首先并行实现将图像上的全日面从日心(对应上面所求重心)位置平移到图像中心位置，即得到中心化图像。中心化过后，再并行实现用双线性插值法将图像转换至极坐标系下。紧接着对极坐标图像角度的 4 个等间距区域分别进行排序，排序方法是并行实现的双调排序算法。从排完序的极坐标图像中直接取出所需的 4 条临边昏暗曲线，将两条曲线用现有函数进行点积、求和等运算，再将结果代入公式即可得两条曲线的相关系数。利用比较所得最小相关系数如判断有云污染，则不需要修复，否则继续下面去除云污染的步骤。

在使用模板法去除云污染时，要用到模板的标准临边昏暗轮廓。生成模板算法与检测云算法很相似，图 7(a)展示了 Hα 全日面像生成模板系统的界面。先通过并行相除得到带有太阳活动细节的云层透射率。紧接着利用 CUFFT 库中函数进行傅里叶变换将图像转换到频域。再并行生成一个巴特沃斯低通滤波器，将其与频域图像并行相乘，即可得到滤除高频部分的频域图像。然后将滤除高频部分的频域图像利用傅里叶逆变换转回时域得到消除太阳活动细节的云层透射率。最后将中心化云污染像与云层透射率并行相除得到最终去除云污染的图像。

对于系统中一些重要参数，根据以往的数据实验统计设置默认值。为了让用户可以在合理范围内进

行调整，同时将这些参数框放在界面上。如图 7(b)为 Hα 全日面像云污染实时识别和修复系统界面。

CPU
开始
打开实时处理界面
检测是否有新图像传入
否
显示前一幅图像处理结果
是
读入图像数据并显示在界面上
图像数据
将处理结果显示在界面上
处理结果

GPU
图像数据二值化
根据二值化图像求出椭圆长短轴之比
判断太阳区域是否为标准圆
否
生成一张打叉的图片说明云污染严重无法修复
是
将太阳从日心平移到图像中心
图像转换到极坐标系
将极坐标图像排序并求其最小相关系数
根据最小相关系数判断是否有云污染
否
中心化图像作为最终结果
是
去除云

图 6 Hα 全日面云污染实时识别和修复系统

Fig. 6 The flow chart of the real-time processing system for detecting and removing cloud shadows on Hα full-disk solar image

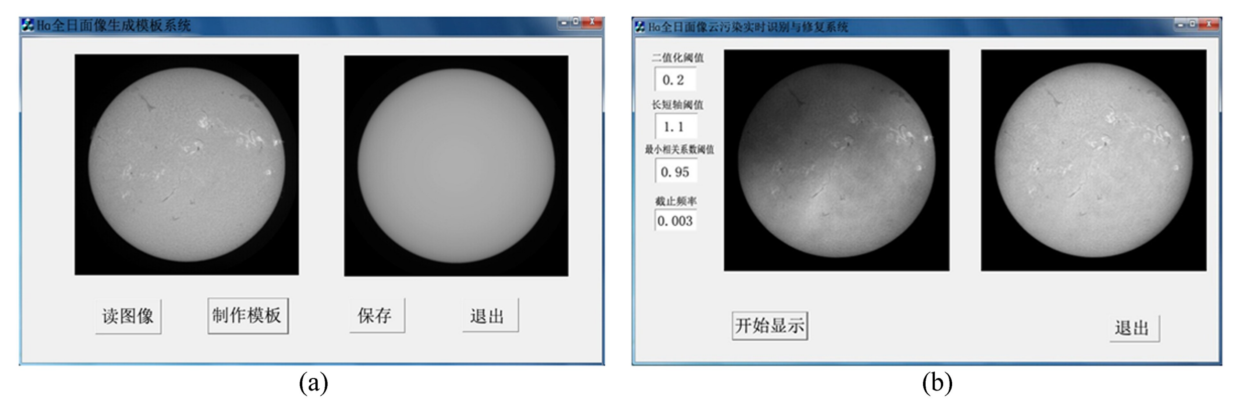


图 7 (a)生成模板系统界面；(b)云污染实时识别和修复系统界面

Fig. 7 (a) The system interface of generating templates; (b) The system interface of real-time detecting and removing cloud shadows

3　系统性能测试

实验采用的计算机配置为：中央处理器为英特尔 Pentium(奔腾)双核 E5200 处理器，图形处理器卡为 Nvidia 公司的 GeForce GT 430。操作系统为 Windows 7 旗舰版 64 位。软件开发环境为 Microsoft Visual Studio 2010 下基于对话框的微软基础类库(Microsoft Foundation Classes，MFC)和 CUDA 6.5。

对系统中各主要运算在图形处理器中花费的时间进行测量，测试数据为图 1(b)云污染图像(尺寸为 2 048×2 048 像素)，得到如表 1 的测试结果。

表 1　系统中各运算在图形处理器中处理时间

Table 1　The processing time of each operation in GPU

运算任务			运算时间/ms	占总时间比/%
将图像原始数据从中央处理器复制到图形处理器			14.2	3.8
重度云污染图像的判断	图像二值化	3.9	33.2	8.8
	求各像素点横纵坐标	5.8		
	求椭圆重心坐标	4.1		
	求椭圆各点与重心偏移量	6.9		
	求椭圆长短轴比值	4.8		
无云污染图像的判断	图像中心化	10.6	48.8	13.0
	图像直角坐标到极坐标转换	8.4		
	极坐标图像行排序	26.8		
	取出 4 条中值临边昏暗曲线	0.3		
	计算 4 条曲线的最小相关系数	1.4		
云污染图像的修复	云污染图像除以模板标准临边昏暗轮廓	9.9	240.7	64.0
	将实数转为复数	2.5		
	傅里叶变换	77.5		
	生成巴特沃斯低通滤波器	45.6		
	低通滤波	3.5		
	逆傅里叶变换	72.3		
	复数转为实数	9.4		
	云污染图像除以云层透射率	9.9		
将处理结果数据从图形处理器复制回中央处理器			21.7	5.7
其他			17.5	4.7
图形处理器处理总时间			376.1	100

经测试，除了图形处理器的处理时间外，中央处理器的总运行时间为 342 ms。其中，中央处理器花费的时间包括：从硬盘读取原始图像数据和模板图像数据(154 ms)、在界面上显示原始图像和修复后的图像(160 ms)、其余运算(28 ms)。所以，该系统处理总时间为 718 ms。

4　讨　论

图形处理器对高密度图像数据的并行处理能力在系统中得到了充分体现。对于一幅 2 048×2 048 像素的云污染 Hα 全日面图像，系统在图形处理器中完成从检测到修复等处理运算的总时间为 376.1 ms。另外，系统在中央处理器中完成读图像、显示图像等处理总耗时为 342 ms。针对常规 1 min 采样间隔的 Hα 全日面观测，可以满足观测过程中对处理结果实时显示的需要。为了发挥图形处理器的并行优势，在修复过程中采用频域巴特沃斯低通滤波的方法。通过使用图像结构相似度算法对修复后图像做

质量评价，发现该方法与前人所用的中值滤波法和多尺度形态学滤波法修复效果相当。系统除了可以处理 GONG 的数据，还使用怀柔太阳观测基地的 Hα 全日面像进行测试，测试结果正常。此外，这个系统不但可以用于 Hα 全日面像的云污染处理，也可用在其他全日面像的处理中。

虽然统一计算设备架构的 Thrust 函数库提供了对于任意长度数据的排序函数，但需要先将数据放入向量容器，且无法实现数据集间的并行，导致排序时间过长(单个数据集耗时 0.5 ms，总耗时 1.7 s)。而采用双调排序不但数据集内部的排序效率高于库中的函数，而且可以实现数据集之间并行(受共享显存的限制，单次可并行数据集有限，经测试每次并行 12 个数据集总耗时最短，为 26.8 ms)。虽然该算法对所排数据的长度有约束(长度必须为 2 的幂，因此将极坐标角度设为 512)，但并不影响后续的判断。

在云污染的修复环节，生成巴特沃斯低通滤波器用时 45.6 ms，相对于除傅里叶正反变换的其他运算来说，这是一个十分耗时的运算。主要原因是每个线程所承担的运算量较大，而图形处理器单个处理单元性能不及中央处理器内核。曾尝试将该滤波器做成一张图像保存在硬盘中，让系统每次从硬盘中读取然后复制到图形处理器中参与运算。但测试发现，读取用时 77 ms，再复制到图形处理器用时 14.2 ms，总耗时约为在图形处理器中生成的两倍。

然而，系统也存在一些不足之处。首先是检测可修复云污染算法，针对云污染对称均匀分布的情况检测失效(当然这种情况极少出现)。其次是频域滤波的缺陷：(1)当云污染比较严重时，如果通过增大滤波器的截止频率更大限度地去除云污染，则同时会去除很多太阳活动细节；(2)图像在太阳边缘处存在较大梯度差使得滤波时出现一些边缘效应。

要强调系统的主要目的是在观测时实时显示云污染修复后的太阳图像，因而着重关注处理速度。虽然对修正后的数据也进行了保存，但如果需要进一步的对这些云污染图像进行高精度处理，如对各种太阳活动现象进行自动识别、分割和测量，采用文[5]提出的多尺度形态学滤波更为合适。如果将多尺度形态学滤波并行化作为后续需要考虑的工作，同时也需要考虑对由于不精确的平场改正、快门效应以及图像饱和等原因所致畸变的 Hα 全日面图像的修复，使得系统有更广的适用范围。进一步验证系统涉及的云污染识别与修复原理是否同样适用于太阳局部像的云污染处理，如果适用再利用图形处理器技术开发相应的处理软件。

致谢：同时感谢由 NOS 主持的 GONG 项目和国家天文台怀柔观测基地所提供的 Hα 全日面图像。

参考文献：

[1] The GONG network of sites [EB/OL]. [2015-06-09]. http://gong.nso.edu/sites/.

[2] Hill F, Fischer G, Forgach S, et al. The Global Oscillation Network Group site survey, 2: results [J]. Solar Physics, 1994, 152(2): 351-379.

[3] Fuller N, Aboudarham J, Bentley R D. Filament recognition and image cleaning on Meudon Hα spectroheliograms [J]. Solar Physics, 2005, 227(1): 61-73.

[4] 罗鹭鸶，杨云飞，季凯帆，等. 基于临边昏暗轮廓的 Hα 全日面太阳图像质量评价 [J]. 天文学报, 2013, 54(4): 392-400.
Luo Lusi, Yang Yunfei, Ji Kaifan, et al. The quality assessment of Hα full disk solar image based on limb darkening profile [J]. Acta Astronomica Sinica, 2013, 54(4): 392-400.

[5] Feng Song, Lin Jiaben, Yang Yunfei, et al. Automated detecting and removing cloud shadows in full-disk solar images [J]. New Astronomy, 2014, 32(10): 24-30.

[6] 杨云飞，冯松，季凯帆. 畸变全日面观测图像的修复 [J]. 天文研究与技术——国家天文台台刊, 2014, 11(1): 54-59.
Yang Yunfei, Feng Song, Ji Kaifan. A new approach for correcting distortions in full-disk solar images [J]. Astronomical Research and Technology——Publications of National Astronomical

Observatories of China, 2014, 11(1): 54-59.

[7] 朱海波，杨云飞，邓辉，等. Hα 全日面太阳图像云污染的去除 [J]. 天文研究与技术——国家天文台台刊，2014，11(2)：134-139.

Zhu Haibo, Yang Yunfei, Deng Hui, et al. Removal of cloud extinction for Hα full-disk solar image [J]. Astronomical Research and Technology——Publications of National Astronomical Observatories of China, 2014, 11(2): 134-139.

[8] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.

[9] Goyal N, Ormont J, Sankaralingam K, et al. January 2008 signature matching in network processing using SIMD/GPU architectures [R]. Cs Technical Reports, 2008.

[10] Kirk D. NVIDIA CUDA software and GPU parallel computing architecture [C] // Proceedings of the 6th International Symposium on Memory Management. 2007: 578.

[11] Shen Yangbin, Lin Jiaben, Ji Kaifan, et al. New real-time correlation solar observing system based on GPU for acquiring the deep-integration magnetogram [J]. New Astronomy, 2013, 25(2): 32-37.

[12] 施正，向永源，邓辉，等. 一米真空太阳望远镜 Level1 级图像选帧的 GPU 实现 [J]. 科学通报，2015，60(15)：1408-1413.

Shi Zheng, Xiang Yongyuan, Deng Hui, et al. A method of Level 1 frames-selection based on GPU for new vacuum solar telescope [J]. Chinese Science Bulletin, 2015, 60(15): 1408-1413.

A Real-Time Image Processing System for Detecting and Removing Cloud Shadows on Hα Full-Disk Solar

Zhang Nengwei, Yang Yunfei, Li Ranyang, Ji Kaifan

(Computer Technology Application Key Laboratory of Yunnan Province, Kunming University of Science and Technology, Kunming 650500, China, Email: flyingwithcloud@sina.com)

Abstract: Sky clouds lead to Hα full-disk solar observed images covered with shadows, which obscure the details of solar features. In this paper, a real-time image restoring system is introduced. For detecting and removing cloud shadows, and displaying the restoring image in real-time, Graphic Processing Unit (GPU) with Compute Unified Device Architecture (CUDA) parallel programming environment is employed to accelerate computing in the system. The parallel computing steps include: (1) identifying heavy cloud covered images by calculating the ratio of major and minor axes of a fitted ellipse; (2) recognizing recoverable cloud-covered images by calculating the symmetric of limb darkening curve; (3) removing cloud shadows by Butterworth low-pass filter in frequency domain. The total processing time is about 0.7s for one cloud-covered image; it meets the need of displaying in real-time. The processing time of each operation taken in GPU is measured. The results show that Fourier transform, inverse Fourier transform and Butterworth low-pass filtering are the most time-consuming part and together they take up 52.9% of the total time. This paper also evaluate the quality of restored image, and the conclusion is that the algorithm can effectively eliminate cloud shadows; meanwhile, it has little effect on the details of solar features. The existing problem and probable improvement are also discussed at the end of this paper.

Key words: Cloud shadows; Parallel computing; Detecting and removing; Butterworth low-pass filtering